



# Stratospheric aerosol injection research and existential risk

John Halstead

58 Geldeston Road, London, E5 8SB, United Kingdom



## ARTICLE INFO

### Keywords:

Stratospheric aerosol injection  
Existential risk  
Climate change  
Geoengineering  
Moral hazard

## ABSTRACT

In the wake of the continued failure to mitigate greenhouse gases, researchers have explored the possibility of injecting aerosols into the stratosphere in order to cool global temperatures. This paper discusses whether Stratospheric Aerosol Injection (SAI) should be researched, on the controversial ethical assumption that reducing existential risk is overwhelmingly morally important. On the one hand, SAI could eliminate the environmental existential risks of climate change (arguably around a 1% chance of catastrophe), and reduce the risks of interstate conflict associated with extreme warming. Moreover, the risks of termination shock and unilateral deployment are overstated. On the other hand, SAI introduces risks of interstate conflict which are very difficult to quantify. Research into these security risks would be valuable, but also risks reducing willingness to mitigate. I conclude that the decision about whether to research SAI is one of 'deep uncertainty' or 'complex cluelessness', but that there is a tentative case for research initially primarily focused on the governance and security aspects of SAI.

Climate change has been known to be a serious problem for decades but as yet very little has been done about it. The prospects of strong action in the future also appear slim, as states have incentives to pass on the costs to other states and to future generations. The Paris Agreement, widely seen to have been a watershed moment in climate negotiations, relies on unprecedented and unrealistic deployment of CO<sub>2</sub> removal technology (Williamson, 2016), and state pledges and promises made in preparation for Paris, even taken at face value, violate the agreed < 2 °C target (Rogelj et al., 2016). With a climate sceptic now US President, pessimism seems even more justified. There is little evidence that there will be a serious course correction in global climate policy in the near future. From the point of view of the future of humanity, this is concerning. Some controversial lines of argument suggest that unless there is a drastic change in climate policy, we are committed to a non-negligible probability of extreme warming, which could arguably threaten advanced human civilisation.

Reducing CO<sub>2</sub> emissions does not have an immediate effect on the level of climate risk. There is a lag of decades between CO<sub>2</sub> emissions and warming, and, unless CO<sub>2</sub> is actively removed from the atmosphere, a large portion of its warming effect will persist for centuries to millennia. It is in this context of political inertia and carbon cycle inertia that, over the last decade in particular, attention has turned to alternative ways to reduce climate risk. Solar Radiation Management (SRM) is a controversial form of geoengineering which would cool the planet by reflecting sunlight back to space. It is the only known way to quickly and relatively cheaply reduce global temperatures. The most researched and discussed form of SRM is Stratospheric Aerosol Injection (SAI). SAI is the sole focus of this paper. Modelling evidence consistently and overwhelmingly indicates that SAI would greatly reduce the environmental risks of climate change. Any reasonable objections to SAI must therefore criticise it on other grounds.

All major reports on geoengineering have called for more geoengineering research, implicitly including SAI research (National Academy of Sciences, 2015; Shepherd, 2009; Schäfer et al., 2015). My point of departure from these reports is ethical: I assume that reducing existential risk is overwhelmingly morally important.<sup>1</sup> The question I then try to answer is: on this highly controversial

E-mail address: [john@founderspledge.com](mailto:john@founderspledge.com).

<sup>1</sup> I explain this idea in more detail in Section 1.

ethical assumption, is further SAI research justified?

Current research suggests that SAI could reduce some risks but introduce others. On the positive side, current research suggests SAI could eliminate the environmental existential risks of climate change, and reduce some of the security risks of climate change. Moreover some risks such as termination shock and unilateral deployment have been overstated. On the negative side, at present it appears that SAI would be very hard to govern and would introduce risks of interstate conflict that are difficult to quantify. These security risks should be the chief concern of funders focused on existential risk reduction.

Usually, increasing knowledge about a technology through research is good, but SAI research could be harmful in two ways:

1. By reducing willingness to mitigate, which all experts agree is required.
2. By being abused by political actors and increasing the risk of irrational and harmful future deployment (which could occur even if SAI does not obstruct mitigation).

In this paper, I argue that if SAI research does not reduce willingness to mitigate, then it would be justified. However, it is extremely unclear whether research will obstruct mitigation *and* whether, if it does, research would be justified. Thus, the decision about whether to research SAI is plausibly one of *deep uncertainty* or *complex cluelessness*.<sup>2</sup> It is extremely unclear whether SAI research would reduce or increase existential risk, and a plausible case can be made that research would have substantial effects in either direction. Noting this deep uncertainty, I argue that research that focuses primarily on the governance and security aspects of SAI would be justified, provided extensive efforts are made to reduce the risk of mitigation obstruction.

## 1. Background on SRM and existential risk

Climate geoengineering – the deliberate attempt to control the Earth’s climate – comes in two forms (Shepherd, 2009):

*Carbon Dioxide Removal* (CDR) – Removing CO<sub>2</sub> from the atmosphere.

*Solar Radiation Management* (SRM) – Controlling global temperatures by reflecting sunlight back to space.

SRM would counteract the warming produced by greenhouse gases by reducing the warming produced by sunlight. SAI is a form of SRM, which would involve the injection of aerosol particles, such as sulphur dioxide, into the stratosphere. SAI could cool the Earth within a matter of months, and would need to be continually replenished for the cooling effect to be sustained.

According to the Open Philanthropy Project, as of 2013, global funding for SRM research was estimated to be around \$11 m per year.<sup>3</sup> For context, in 2014, the US Federal government alone spent around \$2.6bn on climate change research (Leggett, Lattanzio, & Bruner, 2013). No well-documented field experiments involving controlled emissions of stratospheric aerosols have yet been conducted (National Academy of Sciences, 2015, 72).

There is some disagreement about the cost of an SAI programme sufficient to make a significant and sustained reduction in global temperatures (Reynolds, Parker, & Irvine., 2016). Estimates of the annual cost of direct deployment of SAI range from \$1bn to \$10bn (National Academy of Sciences, 2015, 96–97). However, this is likely to be an underestimate both of the direct and total costs of responsible SAI deployment (MacKerron, 2014), which would include monitoring systems (National Academy of Sciences, 2015, 96–97), security to protect the deployment infrastructure (Nightingale & Cairns, 2015), and backup capacity – imperative given the risk of sudden termination (discussed in Section 2.3). Thus, the total annual cost of SAI is likely to be substantially higher than \$10bn, and would probably be closer to \$100bn. This is still orders of magnitude lower than the projected costs of mitigation (IPCC, 2014b, chap. 6).

My focus here is solely SAI’s effects on *existential risk*. I therefore ignore its effects on other important but less severe risks. An existential risk is one that threatens the premature extinction of Earth-originating intelligent life or the permanent and drastic destruction of its potential for desirable future development (Bostrom, 2013). This definition departs from ordinary usage by allowing that an existential risk might not kill everybody. What matters from the point of view of this definition is fundamentally the permanent curtailing of humanity’s potential for desirable future development, however that may come about. A number of philosophers have argued that, due to the sheer number of good lives in the long-run future, an existential catastrophe would be literally astronomically bad, and therefore that reducing existential risk has overwhelming expected value (Parfit, 1984; Bostrom, 2003, 2013). On this view, even the tiniest reduction in existential risk has an expected value greater than the provision of any ‘ordinary’ good, such as saving a billion lives. As a consequence, all of our efforts ought to be focused on reducing existential risk, as these will swamp any benefits that do not affect the long-run future of humanity. These ethical assumptions are of course highly controversial, but I take them for granted here.<sup>4</sup> The main motivation for exploring their assumptions is that they are thought by many leading philosophers to constitute the most plausible theory of how we ought to take the welfare of future generations into account.

With respect to the question of SAI, this implies that we should focus exclusively on the most extreme risks associated with it rather than its less severe consequences. For example, suppose that SAI deployment would slightly increase existential risk, but would greatly decrease the risk of non-existential but catastrophic harms to the current population. On the view under consideration, we nevertheless ought *not* to deploy SAI.<sup>5</sup>

<sup>2</sup> Morrow (2014) and Elliott (2010) reach a similar conclusion. For a discussion of complex cluelessness see (Greaves 2016).

<sup>3</sup> See [http://www.openphilanthropy.org/research/cause-reports/geoengineering#Who\\_else\\_is\\_working\\_on\\_this](http://www.openphilanthropy.org/research/cause-reports/geoengineering#Who_else_is_working_on_this). I do not know of any other more recent estimates of overall funding.

<sup>4</sup> There are weaker less contentious versions of this thesis, which may also have similar implications.

<sup>5</sup> It is worth mentioning that even if reducing existential risk is not overwhelmingly important, the tail risks of climate change still plausibly dominate the overall

The focus on existential risk reduction entails a focus on the highest impact risks, which tend to have low probability. This paper may therefore seem alarmist. However, under the current ethical assumptions, the highest impact risks dominate the risk assessment.

## 2. The environmental benefits and risks of SAI

The benefits and risks of SAI can be divided into *environmental* benefits and risks, and *security* benefits and risks. Environmental benefits and risks are those that stem from how SAI directly interacts with the environment. Security benefits and risks are those that stem from how SAI affects the risk of interstate conflict. Some of the benefits and risks cut across this distinction, but the distinction is nevertheless useful.

In this section, I examine the environmental benefits and risks of SAI.

### 2.1. Climate change and existential risk

Before discussing the evidence on SAI, we first need to understand the magnitude of the existential risk posed by climate change. We can roughly model the magnitude of the risk as a product of:

- The probability of existential catastrophe conditional on rates or magnitudes of warming.
- The probability of different greenhouse gas concentration pathways.
- The probability of a rate or magnitude of warming conditional on concentration pathways.

There is significant uncertainty about all these factors and therefore about the overall magnitude of the risk. Here I will use the best available information to arrive at a ‘best guess’ of the probability of catastrophic warming, noting the significant epistemic qualifications this involves. I discuss the security risks of climate change in Section 3.

#### 2.1.1. Warming required for existential catastrophe

It is very difficult to know what level of warming would threaten the long-term future of humanity. Humanity has in the past survived extreme ( $\sim 6^\circ\text{C}$ ) and abrupt regional and global warming (Alley et al., 2003) at a much lower level of technological sophistication to today. However, these warming events were against the background of much colder temperatures to today. Warming of  $> 6^\circ\text{C}$  would bring temperatures that are far outside the experience of *homo sapiens sapiens* and its immediate ancestors, constituting a radical departure from the Holocene climate in which human civilisation has thrived (Baum, 2014, 5). The last time global average temperatures were  $10^\circ\text{C}$  above present values was in the Eocene epoch around 34–55 million years ago (Weitzman, 2011a, 2011b, 281–82). It is inherently uncertain how human civilisation would respond to warming of this magnitude.

One research paper has suggested that warming of  $11\text{--}12^\circ\text{C}$  above pre-industrial levels would render most of the planet uninhabitable due to heat stress alone, even once feasible adaptation measures had been taken (Sherwood & Huber, 2010). Warming of  $> 6^\circ\text{C}$  would render most of the tropics uninhabitable (King, Schrag, Dadi, Ye, & Ghosh, 2015, pt. 2), and would have potentially devastating effects on agriculture:

“The existence of critical climatic thresholds and evidence of nonlinear responses of staple crop yields to temperature and rainfall... thus suggest that there may be a threshold of global warming beyond which current agricultural practices can no longer support large human civilizations... However, current models to estimate the human health consequences of climate impaired food yields at higher global temperatures generally incorporate neither critical thresholds nor nonlinear response functions... Extrapolation from current models nevertheless suggests that the global risk to food security becomes very severe under an increase of  $4^\circ\text{C}$  to  $6^\circ\text{C}$  or higher in global mean temperature.” (IPCC, 2014a, 736)

The magnitude of warming is not the sole determinant of the long-term prospects of human civilisation: the *rate* of warming also matters. Extreme warming would probably take several centuries to occur, as the oceans would first have to absorb a large proportion of the heat generated (IPCC, 2013, 1102–3). The existential risk of slow extreme climate change is plausibly low because we would have more time to adapt and to take action against climate change, and because there would be more time for another risk, such as nuclear war, to kill us off before climate change bites. However, as discussed below, there is a non-negligible chance of abrupt or non-linear changes in the climate system. These provide less scope for adaptation and consequently pose a more serious threat to human civilisation.

There are at present no quantified estimates of the probability of extreme *and abrupt* warming, though there are estimates of the probability of extreme warming. Thus, in what follows, as a simplifying assumption, I focus only on the magnitude of warming. As I argue below, this biases the estimate of existential risk upwards. In this paper, I follow Weitzman (2011a, 2011b) in using warming of  $10^\circ\text{C}$  as an ‘illustrator threshold’ for existential catastrophe. There is significant room for reasonable disagreement about the justifiability of this threshold.

(footnote continued)  
costs of climate change (Wagner & Weitzman, 2015).

### 2.1.2. The probability of different levels of eventual greenhouse gas concentrations

There is significant inertia in the climate system in two important respects: there is roughly a 25 to 50 year lag between CO<sub>2</sub> emissions and eventual warming (Hansen et al., 2005),<sup>6</sup> and it is expected that around 40% of the peak concentration of greenhouse gas emissions will remain in the atmosphere 1000 years after the peak is reached (Solomon, Plattner, Knutti, & Friedlingstein., 2009). Thus, any plausible near-term cuts in CO<sub>2</sub> would have a delayed effect on temperatures. Even if CO<sub>2</sub> emission rates decline over time, greenhouse gas concentrations would continue to increase, and so, consequently, would temperatures.

Which concentration pathway we follow depends on the extent to which the international community can overcome the collective action problems associated with climate change. Greenhouse gas mitigation is a global and intergenerational public good, and there is large uncertainty about when we reach the threshold for dangerous greenhouse gas concentrations (Barrett & Dannenberg, 2012; Nordhaus, 2015; Wagner & Weitzman, 2015). All of these factors produce a severe free rider and first mover problem such that each state has an incentive not to take action, and the current generation has an incentive to free ride on future generations. Moreover, a large portion of the financial costs of mitigation are borne by politically powerful special interests, making adequate regulation even more difficult. Furthermore, all historical analogues for significant emissions reductions have occurred in developed countries (Rogelj et al., 2016). Achieving similar results in developing countries with growing energy-intensive sectors, less capacity, and a greater humanitarian need for energy growth will arguably be more challenging. Finally, climate scepticism is popular in some major emitting countries (Funk & Kennedy, 2016), which makes adequate climate policy less likely. For these reasons, even though climate change has been known to be a problem for decades, efforts to limit greenhouse gas emissions have been grossly unsuccessful.

With the background set, we can now sketch a ‘best guess’ of the probability of different concentration pathways.<sup>7</sup> The unit of ‘carbon dioxide equivalent’ (CO<sub>2</sub>e) describes how much global warming over 100 years a given type of greenhouse gas, such as methane or nitrous oxide, would cause using the functionally equivalent amount of CO<sub>2</sub>. The probability of concentrations of < 500 ppm of CO<sub>2</sub>e by 2100 is, in my view, plausibly < ~5%. This requires that greenhouse gas emissions peak roughly in the next five years (van Vuuren et al., 2011), which would require the imposition of a global carbon price of at least \$20 in the next five years (Rogelj, McCollum, Reisinger, Meinshausen, & Riahi., 2013). This seems unlikely due to the political barriers to climate action. Most low emissions pathways also require the massive deployment of CDR technology in the second half of the century, which is unproven at scale and seems infeasible due to land use and other issues (Smith et al., 2015; Williamson, 2016).<sup>8</sup>

The probability of concentrations > 1000 ppm of CO<sub>2</sub>e by 2100 is greater but still relatively small, in my view plausibly < ~10%. For instance, the high emissions scenario considered by the IPCC – resulting in around 1300 ppm CO<sub>2</sub>e by 2100 – assumes a massive and improbable increase in the share of power produced by coal and no greenhouse gas mitigation policies until the end of the century (Riahi et al., 2011).<sup>9</sup>

If the foregoing assessment is roughly correct, then the vast majority of the probability is distributed between 500 ppm and 1000 ppm. A number of sources suggest that without a significant policy course change, a medium-high pathway is most likely (King et al., 2015, pt. 2), resulting in greenhouse gas concentrations of 800 ppm by 2100. The International Energy Agency takes emissions reduction pledges at face value and predicts that CO<sub>2</sub>e concentrations will be around 700 ppm by 2100 (Wagner & Weitzman, 2015, 31).

It is impossible to be confident in any probability estimate across greenhouse gas concentrations. Nevertheless, using the foregoing background information, we can attempt to give a reasonable quantification of the probabilities, which helps to roughly bound the scale of the risk posed by climate change. A reasonable best guess is presented in Table 1, making the simplifying assumption that concentrations come in increments of 100 ppm of CO<sub>2</sub>e.<sup>10</sup>

The probabilities in Table 1 do not equal 100% because I exclude the tail at the end of the probability distribution (> 900 ppm).

### 2.1.3. The probability of warming conditional on greenhouse gas concentrations

Equilibrium climate sensitivity (ECS) is defined as the equilibrium global surface temperature change following a doubling of atmospheric CO<sub>2</sub> concentration. There is uncertainty about ECS. Current climate models produce a broad range of estimates of ECS, mainly due to difficulties in modelling the radiative effects of clouds (IPCC, 2013, chap. 9 and 12; Stevens & Bony, 2013). If ECS is > 6 °C, then the risk of very bad outcomes is severe even on lower emissions scenarios. According to the IPCC, there is between a 0% and 10% probability that ECS is greater than 6 °C (IPCC, 2013, 16).

The IPCC’s definition of ECS does not take account of a number of important feedbacks, which could, with low probability, produce extreme warming, such as the release of CO<sub>2</sub> and methane from permafrost, and the release of vast quantities of methane from clathrates (IPCC, 2013, chap. 6 and 12). Greenhouse gas concentrations are now higher than they have been for hundreds of thousands of years (IPCC, 2013, 11), bringing significant unknown unknowns. There is great uncertainty and disagreement about the potential for abrupt and nonlinear changes in the climate system, and such changes have occurred in the past in response to

<sup>6</sup> There is uncertainty and disagreement about the exact extent of the lag, with estimates ranging from a decade to a century or longer (Hansen et al., 2005).

<sup>7</sup> In what follows, I assume that atmospheric concentrations of greenhouse gases in 2100 are all that matter for warming. This is a simplification as it also matters whether concentrations are rising or falling at that point in time.

<sup>8</sup> I note a caveat to this below.

<sup>9</sup> It is true that we are currently tracking slightly above the highest emissions scenario (Sanford, Frumhoff, Luers, & Gullette., 2014), but this does not suggest that we will continue to do so for the next 100 years. Trends in coal use seem unlikely to follow the trajectory implied by RCP8.5 (International Energy Agency, 2016).

<sup>10</sup> The increments can be taken to capture the probability of concentrations 50 ppm either side of each increment. This assumption is due to lack of data on warming conditional on concentrations. In principle, we would need the probability density function for greenhouse gas concentrations and for extreme warming conditional on concentrations, and would take the integral over levels of concentrations of  $f(G)*f(T|G)$  (where  $f$  is a probability density).

**Table 1**A best guess of probabilities of eventual greenhouse gas concentrations ( $G$ ) measured in ppm of CO<sub>2</sub>e by 2100.

$G$	400	500	600	700	800	900
$p(G)$	1%	5%	20%	30%	20%	15%

comparable radiative forcing to that currently being produced by humanity (IPCC, 2013, chap. 12; Alley et al., 2003).

Weitzman (2009b) has argued for the use of a conception of generalised climate sensitivity that accounts for important heat-induced feedbacks and weakening of carbon sinks. This conception of climate sensitivity is closer to the idea of *Earth system sensitivity*, which includes slow feedbacks relevant over the timescale of centuries. Due to these positive feedbacks, Earth system sensitivity is likely to be higher than ECS (IPCC, 2013, 82–85). Weitzman argues, using very rough calculations, that the probability distribution across climate sensitivity, thus understood, has a very long and fat tail (Weitzman, 2009b, 2009a, 2011a, 2011b). He suggests the following probabilities of warming conditional on concentrations (Weitzman, 2011a, 2011)<sup>11</sup>:

Using the simplifying assumption that concentrations will come in discrete increments of 100 ppm, the unconditional probability of existential catastrophe level warming is as follows. Let  $G_i$  be atmospheric greenhouse gas concentrations in discrete increments of 100 ppm of CO<sub>2</sub>e, such that  $G_1 = 100$  ppm,  $G_2 = 200$  ppm, ...  $G_n = n \times 100$  ppm; and let  $T$  denote warming in excess of 10 °C.

$$p(T) = \sum_{i=1}^n p(G_i)p(T|G_i)$$

Deducing from the estimates in Tables 1 and 2, the unconditional probability of existential catastrophe-level warming is ~3.5%.

I use Weitzman's estimate of climate sensitivity because it attempts to account for climate feedbacks which are important from the point of view of existential risk reduction. However, Weitzman's ECS estimate is highly controversial, and there are a few reasons to think it may be too high. Nordhaus (2011a, 2011b) has criticised Weitzman's analysis of the sample of IPCC model probability distributions across ECS. Weitzman (2009a) has defended his approach and noted that even if Nordhaus' approach is correct, the probabilities in Table 2 would be reduced by around 60%, which still suggests that the risk of existential catastrophe is ~1.5%.

Another question concerns the IPCC estimates of ECS which form the basis of Weitzman's estimates. Probability distributions across ECS are highly sensitive to the shape of the Bayesian prior used, and there is significant expert disagreement about what this should be (IPCC, 2013, 921–26; Annan & Hargreaves, 2011). There is also conflict between estimates of ECS based on instrumental data – which first produced observations in the mid-19th Century (IPCC, 2013, 4) – and paleoclimatic data. A number of models using instrumental observations suggest that the probability of ECS > 10 °C is non-negligible (IPCC, 2013, 925), whereas paleoclimatic data suggests that ECS of that magnitude cannot be reconciled with the data (IPCC, 2013, chap. 10).

Finally, and perhaps most importantly, humanity would be able to respond if climate sensitivity turns out to be high, provided that runaway feedback loops were not already underway.<sup>12</sup> If atmospheric concentrations hit around 900 ppm CO<sub>2</sub>e and we had strong evidence that warming will exceed 10 °C, the international community would have very strong incentives to decarbonise rapidly and deploy CDR on a very large scale,<sup>13</sup> which would reduce greenhouse gas concentrations and diminish the existential threat. In short, it is implausible that humanity would sleepwalk into catastrophe if it still had the means to prevent it.<sup>14</sup> Nonetheless, such measures would be ineffective against rapid and extreme warming, which suggests that this kind of warming accounts for the vast majority of the existential risk of climate change. There are at present no quantified estimates of this particular kind of warming, but it must be much lower than the figure implied by the estimate of climate sensitivity above because that accounts warming of all speeds.

Thus, the headline estimate I have produced in this section is highly controversial and some lines of argument suggest that the existential risks of climate change are (much) lower, plausibly < 1%. This controversy should be borne in mind in what follows.

If the environmental risks of climate change are much lower, it is unclear (on the current ethical assumptions) whether this strengthens or weakens the case for SAI research. The desirability of SAI research would be determined by whether it reduced or increased the risks of catastrophic political conflict, and, as I will discuss below, its impact on such risks is unclear. However, the existential risk-driven case for SAI research rests in part on its ability to eliminate the environmental existential risks of climate change. If SAI would not have these benefits, then funders concerned with existential risk should to some extent deprioritise it relative to other cause areas, such as biosecurity or nuclear war.

<sup>11</sup> It should be noted that uncertainty about carbon and methane feedbacks is actually uncertainty about eventual greenhouse gas concentrations, rather than about the sensitivity of the climate to given greenhouse gas concentrations. The estimates of concentrations in Table 2 can be thought of as the product of direct anthropogenic emissions, with the probabilities in the second row expressing the probability of warming eventually caused by these initial concentrations, with carbon and methane feedbacks and weakening carbon sinks as an intermediate cause.

<sup>12</sup> I am grateful to an anonymous reviewer for pushing me on this point.

<sup>13</sup> At moderate levels of warming, as discussed in Section 5, the social and environmental barriers to very large scale CDR would plausibly be prohibitive. However, if humanity were faced with an existential catastrophe, the social and other environmental objections to CDR would lose much of their force.

<sup>14</sup> I am grateful to Andrew Snyder-Beattie for suggesting this idea.



**Table 2**

Probabilities of eventual warming of  $> 10^\circ\text{C}$  above pre-industrial level for given  $G$  measured in ppm of  $\text{CO}_2\text{e}$ ,  $T$  is the probability of warming  $> 10^\circ\text{C}$ .

$G$	400	500	600	700	800	900
$p(T   G)$	0.2%	0.83%	1.9%	3.2%	4.8%	6.6%

## 2.2. SAI, climate change, and existential risk

When discussing the effects of SAI deployment, it is useful to distinguish three worlds: a Baseline Planet – the world as it is today; a Greenhouse Planet in which greenhouse gas emissions continue as expected over the next hundred years but there is no climate engineering; and an Engineered Planet, which is like the greenhouse world except that SAI is also deployed (Morton, 2015, chap. 4). The comparison between the Greenhouse Planet and the Engineered Planet is the relevant one for our purposes.

Most modelling studies of SAI have thus far assessed its impacts using its effects on temperature and precipitation as a proxy. It is important to note that the effects of SAI depend on how it is done, and this would be a matter of political choice. Modelling studies thus far suggest that SAI cannot completely reverse all the effects of greenhouse gases, as either temperature or precipitation could be returned to earlier levels, but not both at the same time. Thus, the Engineered Planet would probably be worse than the Baseline Planet. However, modelling studies have consistently shown that the Engineered Planet would have reduced temperature and precipitation anomalies in most or all regions, relative to the Greenhouse Planet (Kravitz et al., 2014; Caldeira, Bala, & Long, 2013; Irvine, Kravitz, Lawrence, & Muri., 2016; Schäfer et al., 2015, 49–54; Keith & Irvine, 2016). If used to reduce the rate and magnitude of warming, SAI could also reduce the risk of damaging positive feedbacks and tipping points in the Engineered Planet relative to the Greenhouse Planet (Irvine, Schäfer, & Lawrence, 2014).

Evidence of the impact of SAI on natural and human systems such as agriculture, health, water resources, and ecosystems is limited. Studies of SAI's effects on overall water availability are notably absent, for example. The handful of studies on agricultural output show broadly positive results relative both to a Greenhouse Planet and, in some studies, to a Baseline Planet, with SAI increasing yields in part aided by the  $\text{CO}_2$  fertilisation effect (Irvine et al., 2017).

If SAI is to be a useful tool in reducing the *existential* risks of climate change, it would have to reduce and delay positive feedbacks, or partially offset very high greenhouse gas radiative forcing. Models suggest that it becomes harder to keep temperature and precipitation within safe bounds for all regions as greenhouse gas radiative forcing increases (see for example Ricke, Granger Morgan, & Allen, 2010). This said, on the current ethical assumptions, regional damage would be justifiable provided SAI reduced global existential risk overall. Of course, the accuracy of existing models is highly uncertain, especially with regard to precipitation, so all these results should be interpreted with some caution.

SAI is a unique climate policy tool because it works within a matter of months, whereas  $\text{CO}_2$  mitigation takes decades to have an effect on temperature. There is disagreement about the conditions in which SAI should be deployed. In the early debates over SAI, some argued SAI should be used only in a 'climate emergency' (Shepherd, 2009). However, the two leading proposals in the current debate are that SAI should be used well in advance of a climate emergency, firstly to slow the rate of warming, and/or secondly to reduce the peak level of warming as we decarbonise the economy and eventually reduce greenhouse gas concentrations using CDR (Keith, 2013). (However, note that there is a distinction between the situation in which SAI ideally *should* be deployed and when it would be *most likely* to be deployed. It is plausible that, given the governance challenges involved with SAI, SAI would be unworkable until climate change had become particularly damaging for the vast majority of states).

Whatever the conditions in which SAI should be deployed, the evidence thus far does suggest that SAI is a potentially useful tool specifically with regard to the environmental risks of climate change. SAI could be used to offset a significant portion of warming of  $> 10^\circ\text{C}$ , and thereby eliminate the environmental existential risks of climate change, thereby reducing existential risk by  $< 1\%$  to  $3.5\%$ . Because the greenhouse gas forcing would be so strong, SAI would bring severe regional costs, but the benefits, in terms of existential risk reduction, dominate.

## 2.3. Termination shock

Due to the long-term warming effects of  $\text{CO}_2$  and to inertia in the carbon cycle, to provide significant benefits, SAI would probably have to be deployed for around 100 years or more. One of the main risks of SAI stems from the fact that it could be terminated suddenly causing rapid and damaging warming. According to Baum, Maher, and Haqq-Misra, (2013, 168), "while the outcomes of the double catastrophe are difficult to predict, plausible worst-case scenarios include human extinction". There are some reasons to believe that the risk of termination shock has often been overstated (Parker & Irvine, 2016).<sup>15</sup>

Firstly, termination shock is only a problem for very thick stratospheric veils. If a more moderate stratospheric veil were used, catastrophic termination shock risk would be reduced (Kosugi, 2013). Secondly, SAI would not necessarily have to be stopped abruptly. SAI could be phased out gradually and thereby reduce the rate at which we arrive at a certain magnitude of warming (Keith & MacMartin, 2015). If we reduce greenhouse gas emissions and use CDR to reduce greenhouse gas concentrations, then SAI could be used to reduce both the rate and absolute magnitude of warming.

<sup>15</sup> The following argument is borrowed entirely from (Parker & Irvine, 2016).

Thirdly, for there to be a termination shock, it must be prohibitive to resume SAI within a buffer period of a few months. Given the costs of sudden termination and the cost of SAI, if SAI were stopped for some reason, then, under normal circumstances, other countries would be both willing and able to step in and resume SAI. Moreover, each country would have strong incentives to build resilience into the system, and even if countries did not do this, it is plausible that countries would be able to develop SAI delivery mechanisms quickly from scratch, given their cost. The two most promising delivery methods appear to be custom-built planes and hoses tethered to balloons (Shepherd, 2009). Therefore, for an event to prevent SAI resumption, it would have to be so catastrophic that it prevents the use of planes or large balloons for several months. Thus, a very specific and severe catastrophe is required. Candidates include a very severe engineered pandemic, a very large nuclear war, a large asteroid strike, and so on. It is extremely uncertain how likely they are (Global Priorities Project, 2017, sec. 1), making it very difficult to assess the overall significance of termination shock risk. Nonetheless, these events all seem to have low probability, which suggests that termination shock risk is small.

The risk of termination shock also interacts with SAI's security risks. Termination shock risk would be exacerbated if SAI deployment reduces willingness to mitigate. If it did, a thicker veil would be needed for a longer period of time.

## 2.4. Unknown environmental benefits and risks

Research on SAI is in its infancy and our understanding of the climate system is very imperfect. Given this, SAI could have currently unforeseen effects. However, SAI has (admittedly imperfect) natural analogues. Volcanic eruptions have in the past produced global cooling by ejecting particles into the stratosphere. For example, the 1991 Mount Pinatubo eruption injected around 20 million tonnes of sulphur dioxide into the stratosphere (National Academy of Sciences, 2015, 7), without coming close to threatening existential catastrophe. For SAI, 1 to 5 million tonnes of sulphur per year would be required (Shepherd, 2009, 32). Although SAI would be over much longer timeframes (decades to centuries), this analogue does suggest that the unknown environmental *existential* risks of SAI are negligible. If SAI were to be deployed, it would be rational to first invest significant amounts into climate modelling. A gradual phase-in would also reduce unknown risks.

## 3. The security benefits and risks of SAI

If there was a single benevolent and effective global political authority, the case for SAI research and, arguably, future deployment would be strong. But this is not the world in which we live, and SAI needs to be assessed in the context of nation-states pursuing their own goals within relatively weak external constraints. SAI creates some unprecedented political challenges that could increase the risk of interstate conflict, but SAI would also bring some security benefits.

### 3.1. Security benefits

If modelling evidence so far is correct, then SAI would reduce climate impacts, which would in turn help to reduce the security risks that are likely to be associated with global warming. It is very difficult to quantify the expected size of the security risks of climate change (IPCC, 2014a, chap. 12). In a  $> 6^{\circ}\text{C}$  world, there is a serious risk of severe agricultural disruption and of unprecedented migration from the tropics. In recent years, much milder disruption has brought serious political instability. Thus, the security risks associated with climate change appear non-negligible.

### 3.2. Security risk: direct military use

The vast majority of states are extremely unlikely to use SAI or weather modification more generally, as weapons of war:

“Such applications are banned by an international convention, have limited military use given the inherent unpredictability of weather systems... and compete against cheaper, more effective alternative means of achieving the same military ends.” (Nightingale & Cairns, 2015, 8)

However, there is one caveat to this. It is possible that certain forms of SAI could be used as a doomsday device (Morton, 2015, 342–43). Keith (2010) has discussed the speculative possibility of self-levitating engineered nanoparticles that could loft above the stratosphere and stay there for around a decade, rather than for one to two years.<sup>16</sup> If particles such as these were deployed for around a century, they could conceivably start an ice age (Morton, 2015, 342).

However, the existential risks posed by the weapon as described appear negligible. The levels of nanoparticles would have to be continually replenished over the course of a century, and consequently other actors would be willing and able to disrupt its deployment, and its effects could be counteracted by counter-geoengineering (discussed below).

This said, further research into SAI could inadvertently increase knowledge about a geoengineering-based doomsday weapon that does not need to be continually replenished and so is less susceptible to disruption. In this way, SAI research could constitute an *information hazard*, or could increase attention to a geoengineering-based weapons and thereby constitute an *attention hazard* (Bostrom, 2011). Research funders should place strict restrictions on research into weaponised forms of SAI. It might also be

<sup>16</sup> This would not be a form of stratospheric aerosol injection, but would be a form of SRM.

worthwhile for concerned funders and governments to conduct research into the possibility of a more threatening geoengineering doomsday weapon. This research should probably be classified.

Overall, the risks posed by direct military use currently appear small relative to the environmental benefits.

### 3.3. Security risk: unilateral deployment

SAI appears to create some major governance challenges, firstly because it could be unilaterally deployed, and secondly because it politicises the weather and its success depends to some extent on continued agreement by affected states over long time periods.

Turning first to unilateralism, some have argued that since SAI is so cheap and works so quickly, a wide array of actors, including a coalition of large states, a single small state or even a rich billionaire could unilaterally deploy SAI (Victor, 2008). SAI is therefore subject to the *unilateralist's curse* or *free driver problem*: because deployment capability is so widely dispersed, if states act in what they believe to be the national or the global interest, SAI deployment will tend to be suboptimal by damaging other regions (Bostrom, Douglas, & Sandberg, 2016; Wagner & Weitzman 2015).<sup>17</sup> Regionally damaging SAI would increase the risk of interstate conflict.

In my view, the risks of unilateralism are overstated. Firstly, as argued in Section 1, the cost estimates relied upon are likely to be a significant underestimate, plausibly by an order of magnitude. Secondly, as Parson (2014, 98) has argued “these scenarios overstate the distribution of capabilities and thus the risk of unilateral action, because they focus too narrowly on financial cost as the determinant of capability and neglect other, non-financial, requirements and constraints.” An SAI programme large enough to make a non-trivial sustained impact on the climate would be hard to conceal and vulnerable to military attack.

“[U]nilaterally achieving a climate alteration that matters would require not just the money, technological capability, and delivery assets, but also the command of territory, global stature, and ability to deploy and project force necessary to protect a continuing operation against opposition from other states, including deterring their threats of stopping it through military action.” (Parson 2014, 99)

This suggests that scenarios in which small states or rich individuals deploy SAI are vanishingly unlikely.

Indeed, Horton (2011) has persuasively argued that SAI is actually characterised by a logic of multilateralism. The success of an actor's SAI programme would depend on whether other actors were also pursuing their own SAI programme and would be ineffective without coordination. Moreover:

“States opposed to geoengineering have a number of tools at their disposal to counteract climate interventions. In the case of SAI, for example, fluorocarbon gases could be deployed to offset cooling effects. Alternatively, the strategic use of black carbon could neutralize artificial albedo enhancement.” (Horton, 2011, 62)

In short, if powerful actors were opposed to an SAI programme by a state or a collective of states, they could effectively discourage it using ordinary military threats or by counteracting the effects of SAI.

The foregoing suggests that the decision to deploy unilaterally would not be taken lightly, given the incentives created by conventional military threats and the ease with which SAI schemes can be disrupted. Even for a case in which a major power is facing very severe climate impacts, SAI without support from other major powers would likely either be counter-productive or ineffective. In my view, this suggests that unilateral deployment even by a powerful state or some coalitions of powerful states is not a serious danger, provided that there are some dissenting major powers (though it should be noted that many experts disagree).

This analysis still allows that the interests of states that lack international standing could be ignored. However, the *existential* risks of this are negligible because this could not precipitate a nuclear war between major powers.

### 3.4. Security risk: politicisation of the weather and unilateral withdrawal

SAI poses some major governance problems, chiefly stemming from the fact that it would produce regionally diverse impacts and would politicise the weather for a long period of time. To bring substantial benefits, SAI would probably have to be deployed for a century or more, which means that robust global agreement on SAI would have to be sustained for that time period. Even if, as the evidence suggests, SAI reduces overall harms relative to a Greenhouse Planet, it is highly likely to cause harm to some countries at some point. One (partial) solution to this is to ensure that those affected are compensated.<sup>18</sup> There is some disagreement about the feasibility of an adequate compensation scheme (Nightingale & Cairns, 2015; Horton, Parker, & Keith, 2014). If the harm is particularly severe – thousands of deaths or more – it may be difficult to find a level of compensation that is politically acceptable to the affected party.

Moreover, some countries will experience adverse weather events while SAI is deployed and will attribute these events to SAI, even if SAI is not in fact responsible. The climate system is highly unpredictable and chaotic, so at best we will be able to attain probabilistic causal attribution of adverse events to SAI (Horton et al., 2014). Trust in the probabilistic models would have to be high for severe adverse weather events not to be blamed on SAI by some members of the public or by political leaders.

If countries do blame severe adverse weather events on SAI, the response is likely to be angry and irrational, perhaps including

<sup>17</sup> Bostrom et al assume that each state acts in the global interest, whereas Weitzman assumes that each state acts in the national interest. It is not clear whether Bostrom at all believe that capabilities for unilateral SAI deployment are widely dispersed. They may just assume this as a means to illustrating the unilateralist's curse.

<sup>18</sup> For discussion of the issues surrounding compensation, see (Svoboda & Irvine, 2014; Wong, Douglas, & Savulescu, 2014).



suspicion about the motives of the controlling coalition (Nightingale & Cairns, 2015). If countries are particularly badly affected they may demand SAI termination or, in the extreme case, retribution. The controlling coalition would then have to manage a slow termination while placating the affected party. If the competing parties are nuclear states, this interstate tension could arguably increase existential risk, though the magnitude of the existential risk posed by nuclear war is somewhat unclear. Furthermore, if SAI reduces willingness to mitigate, SAI would have to be deployed for longer and at greater intensity, which would increase the risk of regional damages and would increase the period in which the weather is politicised.

Importantly, the arguments in this and the previous section suggest that the barriers to effective and sustained deployment of SAI are significant. This makes SAI less useful as a tool against the (arguable) existential risks of climate change. Firstly, SAI is only likely to be deployed once climate change starts to impose substantial costs on most or all regions; and secondly, it will be difficult to sustain agreement on an SAI deployment programme over the course of decades. This suggests that SAI is likely to be used only if climate change turns out to be particularly bad.

### 3.5. Balancing security risks and benefits

Overall, in my view, it is at present extremely difficult to judge whether SAI deployment would reduce or increase the risks of interstate conflict: it is deeply unclear whether the security risks would be more severe in a Greenhouse Planet or an Engineered Planet. Although SAI would reduce the security risks of climate change, it would also require unprecedented multi-decadal global cooperation which would plausibly be regularly threatened by politicisation of the weather.

## 4. SAI research and mitigation obstruction

A common criticism of SAI research is that it could be a ‘moral hazard’. I follow Morrow (2014) in calling this the ‘mitigation obstruction’ argument because moral hazard effects depend on voluntary risk sharing between a risk taker and an insurer, which does not characterise the SAI research worry. A natural first pass at defining the mitigation obstruction worry goes as follows: SAI research leads to a reduction of mitigation because of a premature conviction that SAI provides insurance against climate change (Shepherd, 2009, 39). However, this is not quite correct as a description of the mitigation obstruction worry we ought to be interested in. I use Morrow’s (2014) specification in what follows.

What I call the Pernicious Mitigation Obstruction Argument claims *both* that SAI research decreases willingness to mitigate *and* that this would be worse than what would have happened if there had been no SAI research. Reduced mitigation would be problematic because SAI scientists overwhelmingly agree that if SAI is used, it should be accompanied by aggressive greenhouse gas mitigation (Reynolds et al., 2016, 564).

The argument can be stated more formally in terms of climate policy portfolios. We can represent a portfolio as an ordered quadruple,  $(m, a, c, s)$ , denoting levels of mitigation, adaptation, CDR and SAI, respectively. Let  $(m_0, a_0, c_0, s_0)$  represent whatever portfolio policy-makers would choose in the absence of further SAI research. Call this the ‘baseline portfolio’. Let  $(m_{sa}, a_{sa}, c_{sa}, s_{sa})$  represent whatever portfolio policy-makers would choose if SAI research were to proceed in earnest. The pernicious mitigation obstruction argument involves two very different claims. The first is a prediction, and the second is an evaluative claim about the value of the resulting outcome:

1. *Predictive claim:*  $m_{sa} < m_0$
2. *Evaluative Claim:*  $(m_{sa}, a_{sa}, c_{sa}, s_{sa})$  would yield a worse outcome than  $(m_0, a_0, c_0, s_0)$  by virtue of the effect on  $m_{sa}$ .<sup>19</sup>

How one assesses the evaluative claim depends on one’s moral theory, and here I focus only on the effect SAI research has on existential risk. Thus, the claim I examine here is whether SAI research increases existential risk by reducing willingness to mitigate.

The reason we should be concerned with the pernicious mitigation obstruction argument – rather than an argument that asserts only (1) – is that it sheds light on the all-things-considered desirability of SAI research. Merely learning that (1) is true does not tell us about the desirability of SAI research. For example, suppose that SAI research reduces mitigation by a tiny amount, but makes the SAI part of the portfolio considerably better. In this case, (1) would be true, but (2) false, and SAI research ought to be carried out even though it reduces mitigation.

Finally, it is important to note that the relevant baseline for comparison in our model is what policy would *actually* be absent SAI research, not what policy would *optimally* be absent SAI research. SAI research might reduce interest in mitigation relative to optimal policy – e.g. a > \$20 carbon price – but not relative to actual policy absent SAI research.

## 5. Assessing the pernicious mitigation obstruction argument against SAI research

I will assess each part of the pernicious mitigation obstruction argument in turn. Since many of the factors involved in assessing this argument are also relevant to the assessment of the claim that SAI research would increase existential risk even if it does not obstruct mitigation, I also assess that claim in this section.

<sup>19</sup> My specification of the mitigation obstruction argument departs from Morrow by adding the proviso “by virtue of the effect on  $m_{sa}$ ”.

### 5.1. The predictive claim

It is important to bear in mind that (1) is a prediction about how a host of global political actors will respond to an unprecedented environmental and political problem riven by scientific and political uncertainty. Significant uncertainty about the predictive claim is therefore *prima facie* reasonable.

We first need to consider  $m_0$  – how much mitigation is there likely to be on the baseline portfolio? As I argued in Section 2, there are enormous political barriers to action on climate change which are themselves likely to make mitigation strongly suboptimal in terms of minimising existential risk. These barriers exist even if there is no further SAI research. If SAI research does obstruct mitigation, it will most probably only make a very bad situation even worse.

Now, what evidence do we have on whether  $m_{sa} < m_0$ ? Firstly, empirical evidence on the effects of learning about geoengineering on willingness to mitigate is mixed (Burns et al., 2016). Some studies show that some people's willingness to mitigate decreases, but many other studies show no effect or an overall *increase* in willingness to mitigate. It is very difficult to know how far the findings from these studies will be reflected in the real world political response to SAI. Most importantly, the response is likely to be different following a sustained and slick political or industry campaign.

Secondly, some have argued that the mere knowledge of the possibility of SAI, rather than novel SAI research, explains some of its mitigation obstruction effects (Keith, Parson, & Granger Morgan, 2010). This is true, but it is plausible that the mere knowledge of SAI explains a relatively small portion of the potential mitigation obstruction effects. At present, SAI still receives only a very small portion of the overall attention directed to climate change, and few members of the public have even heard of geoengineering (Mercer, Keith, & Sharp., 2011). If SAI becomes more mainstream, by, for example, being discussed extensively and positively in future IPCC reports,<sup>20</sup> it will inevitably receive greater political attention, increasing the risk of mitigation obstruction.

Thirdly, we need to consider the political mechanisms by which SAI research plausibly could inappropriately obstruct mitigation. There are a number of political pressures and cognitive biases which could lead political actors to respond in an inappropriate way to research (Lin, 2013, 694ff; Morrow, 2014). Overconfidence bias could lead politicians and others to be overoptimistic about the prospects of SAI. In addition, CO<sub>2</sub> mitigation is financially and politically costly and provides no benefits to voters over the course of the five year period in which politicians tend to seek re-election. SAI, on the other hand, is relatively cheap and would provide short-term benefits to voters. Thus, there are incentives for politicians to over-rely on SAI in their climate policy portfolio. Greenhouse gas producing industries also face strong incentives to overstate the extent to which SAI could offset the damage done by greenhouse gases. Indeed, there are a number of examples of academics, politicians and popular writers arguing that SAI obviates the need for mitigation (Hamilton 2013, chap. 7).

Experience with 'clean coal' and CDR suggests that the risk of SAI mitigation obstruction is worth taking seriously. 'Clean coal' involves the use of coal for power generation with Carbon Capture and Storage (CCS). CCS captures CO<sub>2</sub> from large point sources such as fossil fuel plants and deposits it where it does not enter the atmosphere, such as underground. It is widely agreed that CCS will be a hugely important tool in reducing climate risk, especially in decarbonising the industrial sector (IPCC, 2014b, 60). However, the coal industry has spent large amounts of money exaggerating the promise of coal with CCS in order to resist stringent but necessary limits to coal power (Hamilton, 2013, chap. 7; The Washington Times, 2008). Indeed, the idea of 'clean coal' is still used today in attacks on mitigation by the Trump administration (Plumer, 2017). In this case, industry advocacy with the help of political expediency distorted the promise of a potentially useful technology in order to campaign against mitigation. This precedent should raise concerns about the mitigation obstruction risks of SAI research.

The treatment of CDR raises similar concerns. CDR will undoubtedly be an important technology in the fight against climate change, but it has been treated in an uncritical and implausible way in IPCC models. With the 2015 Paris Agreement, the international community agreed to keep warming "well below" 2 °C. Limiting the global temperature rise to 2 °C using the currently leading proposed form of CDR – Bioenergy with Carbon Capture and Storage (BECCS) – would probably require crops to be planted solely for the purpose of CO<sub>2</sub> removal on around one-third of the current total arable land on the planet, or about half the land area of the United States (Williamson 2016, 154). Unless climate change starts to impose very severe costs on all regions, deployment of BECCS on this scale would be very improbable. As Smith et al. (2015) note "there is no [CDR method or combination of CDR methods] currently available that could be implemented to meet the < 2 °C target without significant impact on either land, energy, water, nutrient, albedo or cost" (Smith et al., 2015, 7).

CDR has become a major part of the planned climate policy portfolio even though its costs and benefits have not been comprehensively considered in climate negotiations. Overreliance on CDR implies underreliance on mitigation. Consequently, Anderson and Peters (2016, 183) label CDR a "moral hazard par excellence". Pushing against this argument is that, due to the other barriers to political action, mitigation would have been as weak or nearly as weak even if CDR were treated in a more realistic way. Unfortunately, we do not have access to the counterfactual. The point is that this example suggests that the political mechanism for mitigation obstruction cannot be dismissed.

There are plausible possible futures in which SAI is treated in a similar way: political actors start to see SAI as a feasible option without adequately considering its costs and benefits and, as a result, mitigation is unsatisfactory. The main factor pushing against this is that SAI is an intuitively more novel and frightening prospect than 'clean coal' or CDR, which may reduce the extent it can be publicly appealed to as an excuse for inaction.

There are a number of possible ways for funding organisations and researchers to reduce mitigation obstruction risk (Lin, 2013,

<sup>20</sup> SRM is discussed in the latest IPCC reports, but the treatment is relatively critical (see for example IPCC, 2013, Box TS.7).

707–11; Morrow, 2014, 10–11). For example, researchers could sign a well-publicised declaration about the primary importance of mitigation.

## 5.2. The evaluative claim

In my view, the predictive claim is certainly worth taking seriously. There is nonetheless significant uncertainty about the sign and magnitude of the expected mitigation obstruction effects of SAI research. This creates significant uncertainty about the truth of the evaluative claim: it is difficult to know whether SAI research would be a *pernicious* mitigation obstruction risk without knowing the sign or magnitude of its mitigation obstruction effects.

### 5.2.1. A governance and security-focused research programme would have lower risks of pernicious mitigation obstruction risk than an environmental effects-focused research programme

One robust conclusion we can draw is that a research programme focused *initially* and *primarily* (though not entirely) on the governance and security aspects of SAI would have lower risks of pernicious mitigation obstruction than a research programme focused primarily on the environmental effects of SAI.<sup>21</sup> The main reason for this is that whether SAI research turns out to perniciously obstruct mitigation depends on the outcomes of research. If research is unambiguously positive, and shows that SAI could be governed safely and eliminate the environmental risks of climate change, then even if mitigation is reduced, the improvement in the SAI part of the portfolio would more than compensate. If research is unambiguously negative, and shows that SAI would be dangerous and impossible to govern, then research would take SAI off the table and thereby *increase* incentives to mitigate. If research is more equivocal, then there is more scope for distortion and abuse by political and industry forces. This also suggests that the *presentation* of research by researchers is crucial for managing mitigation obstruction risk.

Security-focused research presents lower mitigation obstruction risk than environmental effects focused research simply because, given research outcomes thus far, it is much more likely to produce negative results.

There are two additional reasons to initially favour security-focused research over environmental effects-focused research. Firstly, the most likely path to harmful future deployment is following a research programme that produces very positive results on the environmental effects of SAI, but neglects the security and governance aspects. Political actors could then unintentionally or otherwise latch on to SAI as a silver ‘bullet solution’ without considering the governance and security challenges. A security-focused programme would reduce this risk by making the governance and security challenges more salient in the public debate.

Secondly, as I have argued above, most of the uncertainty about the existential risks of SAI is driven by uncertainty about the security aspects of SAI. Consequently, reducing this uncertainty through research would have greater expected benefits than reducing uncertainty about the environmental effects.

### 5.2.2. Deep uncertainty and pernicious mitigation obstruction

Granting that an initially governance and security-focused research programme would have lower risks of pernicious obstruction than an environmental effects-focused programme, the next question is: does SAI research create pernicious mitigation obstruction risk?

There are two ways in which it might do so. Suppose that SAI research has strong mitigation obstruction effects ( $m_{sa} < m_o$ ), leading to a ‘Greenhouse+ Planet’, with a much stronger greenhouse effect than the planet we would have had without research – the ‘Greenhouse Planet’. There are two possible futures in which, despite SAI research, Greenhouse+ has higher levels of existential risk than Greenhouse:

1. Even though SAI is used as an excuse not to mitigate, it turns out that the governance obstacles to successful deployment are so severe that it cannot be deployed in Greenhouse+. By obstructing mitigation, SAI research has left us with the greater risks associated with Greenhouse+ without giving us viable tools to mitigate them.
2. A governance agreement can be reached in Greenhouse+ and deployment would be better than no deployment in Greenhouse+. However, due to the strong greenhouse effect, deployment carries high risks of regional damages and therefore of interstate conflict. As a result, the overall level of existential risk is higher in Greenhouse+ than in Greenhouse, even though SAI deployment is preferable to no deployment in Greenhouse+.

In my view, there is *deep uncertainty* about whether research will lead to either of these futures. The uncertainty about these risks is deep in the sense that it is extremely difficult to assign justifiable probabilities to the relevant outcomes. The decision therefore appears to be a case of what Greaves (2016) has called ‘complex cluelessness’: there is a plausible case to be made in either direction and it is unclear how to weigh the competing considerations. In cases of deep uncertainty or complex cluelessness, principles such as ‘minimise existential risk’ or the closely related ‘precautionary principle’ do not provide any clear guidance (Morrow, 2014; Elliott, 2010). The decision about whether to research SAI presents us with a risk–risk scenario: both research and the failure to research could plausibly increase existential risk.

Notwithstanding this deep uncertainty, my own view is that initially and primarily governance and security-focused research

<sup>21</sup> Security-focused research would have to in part be supplemented by environmental effects research so that the nature of the governance challenge would be clear. If the governance and security-focused research turns out to be positive, the primary focus could shift to the environmental effects of SAI.

would be justified. Thus far there has been little work on the governance and security aspects of SAI (Horton & Reynolds, 2016). It is therefore reasonable to expect that this work would be highly valuable on the margin, and would improve the SAI part of the portfolio. As argued in the previous section, the risk of pernicious mitigation obstruction stems from the probability of these events:

1. Equivocal research outcomes.
2. Equivocal research outcomes are used as an excuse not to mitigate.

It is difficult to quantify the probabilities of these events, but I think a plausible case can be made that the probability of 1 is sufficiently low, and that we can take steps to significantly reduce the probability of 2. Indeed, as I have argued, it is unclear what the sign of the mitigation obstruction effects of this kind of research would be. Provided strong provisions are made to reduce mitigation obstruction risk, initially and primarily governance and security-focused arguably presents relatively small mitigation obstruction risks. However, I emphasise the weak epistemic status of this conclusion, noting that it is not justified with well-quantified accounts of the relevant risks (barring the environmental existential risks of climate change).

### 5.3. Could research be harmful without obstructing mitigation?

I noted at the start of this paper that SAI research could be harmful even if it does not obstruct mitigation. This could happen if political actors abuse research and deploy SAI without due consideration of the governance and security challenges. This in turn would increase risks of interstate conflict. In my view, the existential risks of this possibility are negligible for an initially security-focused research programme. There are two possibilities to consider. In both there is no mitigation obstruction.

1. Security-focused research increases the risk that SAI is deployed and badly governed.
2. Security-focused research increases the risk that SAI is not deployed, when deployment would be preferable.

To produce either result, security-focused research would have to be misused in an irrational way by political actors. In both cases, the mechanisms of political irrationality required are implausible. Regarding the first possibility, states would individually and collectively have incentives to govern SAI well if it is deployed, given the risks of interstate conflict involved in poor governance and the multilateral logic entailed by SAI. They would therefore have incentives to be rational in their treatment of SAI governance research. It is also difficult to see what political pathologies could lead to the second possibility. If research showed that deployment would be beneficial and that SAI can be safely governed, it is very difficult to see how that could make deployment less likely: if research showed that deployment would be beneficial, states would individually and collectively have incentives to deploy it.

For these reasons, if security-focused research does not obstruct mitigation, it would be beneficial, in expectation. This provides further support for the view that reducing mitigation obstruction risk should be a priority. One arguable caveat to this is the possibility of weaponised forms of geoengineering, discussed in Section 3.2.

## 6. Would SAI deployment obstruct mitigation?

Sections 4 and 5 concerned whether SAI research would obstruct mitigation in a pernicious way. It is a different question whether SAI deployment would obstruct mitigation in a pernicious way. As I have noted in Sections 2 and 3, if deployment does obstruct mitigation, some of the risks associated with SAI would be exacerbated. This is relevant to SAI research funding decisions.

Many of the considerations I discussed in Sections 4 and 5 apply to the pernicious mitigation obstruction argument against deployment, but deployment also raises some novel issues. Most importantly, the analogy to CCS and CDR is not directly applicable because these are not yet deployed on a large scale. Nonetheless, the analogy may be instructive. If SAI is deployed and is successful, then the pressures which made the prospects of CCS and CDR a mitigation obstruction risk would be brought to bear. Politicians and industry groups would have strong incentives to overstate the extent to which SAI can substitute for mitigation. We have so far failed to solve the free rider problem. Solving it once there is an obvious fast-acting partial remedy for climate change would likely be more difficult.

Parson (2014) has proposed an SAI governance scenario called *Pay to Play Linkage*, which would reduce the mitigation obstruction risks of SAI. This scenario deters free riding by making each state's mitigation performance a condition for its participation in decision-making on SAI:

“Under the dual assumptions that (i) all participating states strongly desire a voice in [SAI] decisions, and (ii) the threat to exclude them from such participation is credible, this approach would address the problem of providing effective incentives for states to accept and meet strong mitigation commitments.” (Parson 2014, 107–8)

Parson notes that whether these assumptions hold depends on a number of conditions (Parson 2014, 108). First, whether states desire a voice in SAI decisions depends on whether their interests are likely to be taken account of by the controlling coalition. As states' interests diverge – a function of regional variation in SAI impacts – the cost of being excluded from decisions increases. Second, SAI which imposed severe costs on the excluded states would probably increase conflict risk to such an extent that the threat of exclusion is not credible. The problem might be eased if the SAI interventions are incremental and moderate, which would attenuate the regional disparity of interests, allowing a balance between the disagreeability of exclusion and the credibility of threatening it.

This latter point suggests that the incentives in Pay to Play Linkage would not be effective if SAI were used to counteract extreme warming. However, in this case, it is plausible that the independent incentives to mitigate would be strong even if SAI is deployed because of the severity of the attendant political and regional environmental costs. Moreover, moderate SAI used against moderate

climate change could have counter-catastrophic benefits by reducing the risk of positive climate feedbacks, making Pay to Play Linkage relevant from the point of view of existential risk reduction.

It is unclear whether SAI deployment makes mitigation obstruction more likely or not. It is plausible that deployment would decrease willingness to mitigate, but it is unclear whether this would be pernicious, in the sense defined above.

## 7. Conclusions

I have argued that SAI would bring the following benefits:

- 1 Eliminating the risk of existential catastrophe-level warming attributable to climate change. On some highly controversial lines of argument, the probability of existential catastrophe due to the direct effects of climate change is 3.5%. A number of arguments suggest this estimate could be too high, and that a more realistic figure is below 1%.
- 2 Reducing the security risks associated with climate change.

SAI would also introduce or increase the following risks:

- 3 Security risks associated with the politicisation of the weather.
- 4 The risk of increasing knowledge about a doomsday weapon.
- 5 Termination shock risk

I have argued that factors 4 and 5 present small risks. However, it is extremely difficult to quantify the magnitude of factors 2 and 3; it is unclear whether the risks of interstate conflict would be more severe in a Greenhouse Planet or an Engineered Planet. This is reflected in the fact that there is no expert consensus on the governability of SAI. I have also noted that research into SAI could be an information hazard or an attention hazard for some kind of geoengineering-based weapon, but the risk of discovering a viable and geopolitically destabilising weapon currently appears small.

I have argued that a research programme that focuses primarily and initially on the governance and security risks of SAI would be preferable to one that focuses primarily on the environmental effects of SAI. However, we are arguably in a state of deep uncertainty or complex cluelessness with respect to the question of whether a security-focused research programme would obstruct mitigation and whether it would increase or reduce existential risk. Nonetheless, in my view, further research is justified provided that extensive efforts are made to reduce mitigation obstruction risk. In light of how hard it is to justify this conclusion with well-justified quantifications of risk, further research into *that* meta-level research question would also be worthwhile.

## Acknowledgements

I am grateful to Sebastian Farquhar, Stefan Schubert, and Owen Cotton-Barratt for comments on an embryonic version of this paper. The audience at the Cambridge Conference on Catastrophic Risk, especially Hugh Hunt and Olaf Corry, also provided very useful feedback. For help with this latest version of the paper, I am indebted to the kind and very helpful feedback of Gernot Wagner, Joshua Horton, Peter Irvine, Jesse Reynolds, Niel Bowerman, David Keith, and Seth Baum. I am also very grateful to two anonymous reviewers whose penetrating criticisms have made this paper significantly better. My biggest debt of gratitude is to Andy Parker whose sustained assistance and guidance has been extremely useful.

## References

- Alley, R. B., Marotzke, J., Nordhaus, W. D., Overpeck, J. T., Peteet, D. M., Pielke, R. A., et al. (2003). Abrupt climate change. *Science*, 299(5615), 2005–2010. <http://dx.doi.org/10.1126/science.1081056>.
- Anderson, K., & Peters, G. (2016). The trouble with negative emissions. *Science*, 354(6309), 182–183. <http://dx.doi.org/10.1126/science.aah4567>.
- Annan, J. D., & Hargreaves, J. C. (2011). On the generation and interpretation of probabilistic estimates of climate sensitivity. *Climatic Change*, 104(3–4), 423–436. <http://dx.doi.org/10.1007/s10584-009-9715-y>.
- Barrett, S., & Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences of the United States of America*, 109(43), 17372–17376. <http://dx.doi.org/10.1073/pnas.1208417109>.
- Baum, S. D., Maher, T. M., & Haqq-Misra, J. (2013). Double catastrophe: Intermittent stratospheric geoengineering induced by societal collapse. *Environment Systems & Decisions*, 33(1), 168–180. <http://dx.doi.org/10.1007/s10669-012-9429-y>.
- Baum, S. D. (2014). The great downside dilemma for risky emerging technologies. *Physica Scripta*, 89(12), 128004. <http://dx.doi.org/10.1088/0031-8949/89/12/128004>.
- Bostrom, N., Douglas, T., & Sandberg, A. (2016). The unilateralist's curse and the case for a principle of conformity. *Social Epistemology*, 30(4), 350–371. <http://dx.doi.org/10.1080/02691728.2015.1108373>.
- Bostrom, N. (2003). Astronomical waste: The opportunity cost of delayed technological development. *Utilitas*, 15(03), 308–314. <http://dx.doi.org/10.1017/S0953820800004076>.
- Bostrom, N. (2011). Information hazards: A typology of potential harms from knowledge. *Review of Contemporary Philosophy*, 10, 44–79.
- Bostrom, N. (2013). Existential risk prevention as global priority. *Global Policy*, 4(1), 15–31. <http://dx.doi.org/10.1111/1758-5899.12002>.
- Burns, E. T., Flegal, J. A., Keith, D. W., Mahajan, A., Tingley, D., & Wagner, G. (2016). What do people think when they think about solar geoengineering? A review of empirical social science literature, and prospects for future research. *Earth's Future*, 4(11), 536–542.
- Caldeira, K., Bala, G., & Long, C. (2013). The science of geoengineering. *Annual Review of Earth and Planetary Sciences*, 41(1), 231–256. <http://dx.doi.org/10.1146/annurev-earth-042711-105548>.
- Elliott, K. (2010). Geoengineering and the precautionary principle. *International Journal of Applied Philosophy*, 24(2), 237–253. <http://dx.doi.org/10.5840/ijap201024221>.



- Funk, C., & Kennedy, B. (2016). The politics of climate. *Pew Research Center: Internet, Science & Tech (blog)*, 2016.
- Global Priorities Project (2017). *Existential risk: Diplomacy and governance*.
- Greaves, H. (2016). Cluelessness. *Proceedings of the Aristotelian Society*, 116(3), 311–339. <http://dx.doi.org/10.1093/arisoc/aow018>.
- Hamilton, C. (2013). *Earthmasters: The dawn of the age of climate engineering*. New Haven, Connecticut: Yale University Press.
- Hansen, J., Nazarenko, L., Ruedy, R., Sato, M., Willis, J., Anthony Del Genio, D. K., et al. (2005). Earth's energy imbalance: Confirmation and implications. *Science*, 308(5727), 1431–1435. <http://dx.doi.org/10.1126/science.1110252>.
- Horton, J. B., & Reynolds, J. L. (2016). The international politics of climate engineering: A review and prospectus for international relations. *International Studies Review*, viv013. <http://dx.doi.org/10.1093/isr/viv013>.
- Horton, J. B., Parker, A., & Keith, D. (2014). Liability for solar geoengineering: Historical precedents, contemporary innovations, and governance possibilities. *New York University Environmental Law Journal*, 22, 225–273.
- Horton, J. B. (2011). Geoengineering and the myth of unilateralism: Pressures and prospects for international cooperation. *Stanford Journal of Law, Science & Policy*, 4, 56–69.
- IPCC (2013). In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. Midgley (Eds.). *Climate change 2013: the physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.
- IPCC (2014a). *Climate change 2014: Impacts, adaptation, and vulnerability: Summary for policymakers*. Cambridge University Press.
- IPCC (2014b). *Climate change 2014: Mitigation of climate change: Working group III contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.
- International Energy Agency (2016). *World energy outlook*.
- Irvine, P. J., Schäfer, S., & Lawrence, M. G. (2014). Solar radiation management could be a game changer. *Nature Climate Change*, 4(10), <http://dx.doi.org/10.1038/nclimate2360> [842–842].
- Irvine, P. J., Kravitz, B., Lawrence, M. G., & Muri, H. (2016). An overview of the earth system science of solar geoengineering. *Wiley Interdisciplinary Reviews: Climate Change*, 7(6), 815–833. <http://dx.doi.org/10.1002/wcc.423>.
- Irvine, P. J., Kravitz, B., Lawrence, M. G., Gerten, D., Caminade, C., Gosling, S. N., et al. (2017). Towards a comprehensive climate impacts assessment of solar geoengineering. *Earth's Future*, 1(January), <http://dx.doi.org/10.1002/2016EF000389> [2016EF000389].
- Keith, D. W., & Irvine, P. J. (2016). Solar geoengineering could substantially reduce climate risks—A research hypothesis for the next decade. *Earth's Future*, 4(11), <http://dx.doi.org/10.1002/2016EF000465> [2016EF000465].
- Keith, D. W., & MacMartin, D. G. (2015). A temporary, moderate and responsive scenario for solar geoengineering. *Nature Climate Change*, 5, 201–206.
- Keith, D. W., Parson, E., & Granger Morgan, M. (2010). Research on global sun block needed now. *Nature*, 463(7280), 426–427. <http://dx.doi.org/10.1038/463426a>.
- Keith, D. W. (2010). Photophoretic levitation of engineered aerosols for geoengineering. *Proceedings of the National Academy of Sciences*, 107(38), 16428–16431. <http://dx.doi.org/10.1073/pnas.1009519107>.
- Keith, D. W. (2013). *A case for climate engineering*. Boston review book. Cambridge, Mass: Cambridge, Mass: The MIT Press.
- King, D., Schrag, D., Dadi, Z., Ye, Q., & Ghosh, A. (2015). *Climate change—a risk assessment*. centre for science policy. University of Cambridge [www.csap.cam.ac.uk/projects/climate-change-risk-assessment/](http://www.csap.cam.ac.uk/projects/climate-change-risk-assessment/).
- Kosugi, T. (2013). Fail-safe solar radiation management geoengineering. *Mitigation and Adaptation Strategies for Global Change*, 18(8), 1141–1166. <http://dx.doi.org/10.1007/s11027-012-9414-2>.
- Kravitz, B., MacMartin, D. G., Robock, A., Rasch, P. J., Ricke, K. L., Cole, J. N. S., et al. (2014). A multi-model assessment of regional climate disparities caused by solar geoengineering. *Environmental Research Letters*, 9(7), 074013. <http://dx.doi.org/10.1088/1748-9326/9/7/074013>.
- Leggett, J., Lattanzio, R., & Bruner, E. (2013). *Federal climate change funding from FY2008 to FY2014*.
- Lin, A. C. (2013). Does geoengineering present a moral hazard. *Ecology LQ*, 40, 673.
- MacKerron, G. (2014). *Costs and economics of geoengineering*. *clim. geog. govern.* Working Paper Series <http://geoengineering-governance-research.org/perch/resources/workingpaper13mackerroncostsandeconomicsofgeoengineering.pdf>.
- Mercer, A. M., Keith, D. W., & Sharp, J. D. (2011). Public understanding of solar radiation management. *Environmental Research Letters*, 6(4), 044006. <http://dx.doi.org/10.1088/1748-9326/6/4/044006>.
- Morrow, D. R. (2014). Ethical aspects of the mitigation obstruction argument against climate engineering research. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 372(2031), 20140062. <http://dx.doi.org/10.1098/rsta.2014.0062>.
- Morton, O. (2015). *The planet remade: How geoengineering could change the world*. London: Granta.
- National Academy of Sciences (2015). *Climate intervention: Reflecting sunlight to cool earth*. Washington, D.C: National Academies Press.
- Nightingale, P., & Cairns, R. (2015). *The security implications of geoengineering: Blame, imposed agreement and the security of critical infrastructure*. *Climate geoengineering governance working paper series*. Tech. Rep. School of Business, Management and Economics, Univ. of Sussex.
- Nordhaus, W. (2011a). The economics of tail events with an application to climate change. *Review of Environmental Economics and Policy*, 5(2), 240–257. <http://dx.doi.org/10.1093/reep/rer004>.
- Nordhaus, W. (2015). Climate clubs: Overcoming free-riding in international climate policy. *The American Economic Review*, 105(4), 1339–1370.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Parker, A., & Irvine, P. (2016). *Termination shock presentation*. Future of Humanity Institute.
- Parson, E. A. (2014). Climate engineering in global climate governance: Implications for participation and linkage. *Transnational Environmental Law*, 3(01), 89–110. <http://dx.doi.org/10.1017/S2047102513000496>.
- Plumer, B. (2017). What 'Clean Coal' is — and isn't. The New York Times. [August 23 2017, sec. Climate] <https://www.nytimes.com/2017/08/23/climate/what-clean-coal-is-and-isnt.html>.
- Reynolds, J. L., Parker, A., & Irvine, P. (2016). Five solar geoengineering tropes that have overstayed their welcome. *Earth's Future*, 4(12), <http://dx.doi.org/10.1002/2016EF000416> [2016EF000416].
- Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., et al. (2011). RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, 109(1–2), 33. <http://dx.doi.org/10.1007/s10584-011-0149-y>.
- Ricke, K. L., Granger Morgan, M., & Allen, M. R. (2010). Regional climate response to solar-Radiation management. *Nature Geoscience*, 3(8), 537–541. <http://dx.doi.org/10.1038/ngeo915>.
- Rogelj, J., McCollum, D. L., Reisinger, A., Meinshausen, M., & Riahi, K. (2013). Probabilistic cost estimates for climate change mitigation. *Nature*, 493(7430), 79–83. <http://dx.doi.org/10.1038/nature11787>.
- Rogelj, J., Michel den Elzen, N. H., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., et al. (2016). Paris agreement climate proposals need a boost to keep warming well below 2°C. *Nature*, 534(7609), 631–639. <http://dx.doi.org/10.1038/nature18307>.
- Sanford, T., Frumhoff, P. C., Luers, A., & Gullette, J. (2014). The climate policy narrative for a dangerously warming world. *Nature Climate Change*, 4(3), 164–166.
- Schäfer, S., Lawrence, M., Stelzer, H., Born, W., Low, S., Aaheim, A., et al. (2015). *The European transdisciplinary assessment of climate engineering (EuTRACE): Removing greenhouse gases from the atmosphere and reflecting sunlight away from earth*. Potsdam: Institute for Advanced Sustainability Studies.
- Shepherd, J. G. (2009). *Geoengineering the climate: Science, governance and uncertainty*. Royal Society.
- Sherwood, S. C., & Huber, M. (2010). An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences*, 107(21), 9552–9555. <http://dx.doi.org/10.1073/pnas.0913352107>.
- Smith, P., Davis, S. J., Creutzig, F., Fuss, S., Minx, J., Gabrielle, B., et al. (2015). Biophysical and economic limits to negative CO<sub>2</sub> emissions. *Nature Climate Change*. <http://dx.doi.org/10.1038/nclimate2870> [advance online publication (December)].
- Solomon, S., Plattner, G.-K., Knutti, R., & Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6), 1704–1709. <http://dx.doi.org/10.1073/pnas.0812721106>.
- Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, 340(6136), 1053–1054. <http://dx.doi.org/10.1126/science.1237554>.

- Svoboda, T., & Irvine, P. (2014). Ethical and technical challenges in compensating for harm due to solar radiation management geoengineering. *Ethics, Policy & Environment*, 17(2), 157–174. <http://dx.doi.org/10.1080/21550085.2014.927962>.
- The Washington Times (2008). *Groups spend millions in 'clean coal' ad war*. The Washington Times [December 25, 2008. <http://www.washingtontimes.com/news/2008/dec/25/groups-spend-millions-in-clean-coal-ad-war/>].
- Victor, D. G. (2008). On the regulation of geoengineering. *Oxford Review of Economic Policy*, 24(2), 322–336. <http://dx.doi.org/10.1093/oxrep/grm018>.
- Wagner, G., & Weitzman, M. L. (2015). *Climate shock: The economic consequences of a hotter planet*. Princeton: Princeton University Press.
- Weitzman, M. L. (2009a). *Reactions to the nordhaus critique*. Harvard University [http://projects.iq.harvard.edu/files/heep/files/dp11\\_weitzman.pdf](http://projects.iq.harvard.edu/files/heep/files/dp11_weitzman.pdf).
- Weitzman, M. L. (2009b). On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics*, 91(1), 1–19. <http://dx.doi.org/10.1162/rest.91.1.1>.
- Weitzman, M. L. (2011a). Fat-Tailed uncertainty in the economics of catastrophic climate change. *Review of Environmental Economics and Policy*, 5(2), 275–292. <http://dx.doi.org/10.1093/reep/rer006>.
- Weitzman, M. L. (2011b). A voting architecture for the governance of free-driver externalities, with application to geoengineering. *The Scandinavian Journal of Economics*, 117(4), 1049–1068. <http://dx.doi.org/10.1111/sjoe.12120>.
- Williamson, P. (2016). Emissions reduction: Scrutinize CO2 removal methods. *Nature*, 530(7589), 153–155. <http://dx.doi.org/10.1038/530153a>.
- Wong, P.-H., Douglas, T., & Savulescu, J. (2014). Compensation for geoengineering harms and no-fault climate change compensation. *Climate Geoengineering Governance Working Paper Series*, 8.
- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., et al. (2011). The representative concentration pathways: An overview. *Climatic Change*, 109(1), 5. <http://dx.doi.org/10.1007/s10584-011-0148-z>.